

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 08:24:10 ; Search time 45.7 Seconds  
(without alignments)  
54.389 Million cell updates/sec

Title: US-09-252-828A-2

Sequence: 1 SWFPVQGPADICQCCNKGDG.....TPSHSRQPHVMSQMSRVS 41

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A\_Geneseq.0601:\*  
2: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
3: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
4: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
5: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
6: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
7: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
8: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
9: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
10: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
11: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
12: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
13: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
14: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
15: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
16: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
17: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
18: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
19: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
20: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
21: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
22: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
23: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	242	100.0	41	20	AAV30214
2	235	97.1	424	21	AAV52988
3	229	94.6	372	21	AAV67287
4	226	93.4	223	15	AAV65552
5	226	93.4	223	20	AAV42483
6	226	93.4	223	20	AAV81820
7	226	93.4	223	21	AAV82218
8	226	93.4	223	21	AAV52692
9	226	93.4	223	21	AAV52183
10	226	93.4	223	21	AAV52987
11	225	93.0	372	13	AAV22239

12	192	79.3	424	15	AAV53498
13	151	62.4	426	15	AAV55199
14	151	62.4	426	20	AAV42472
15	151	62.4	426	20	AAV81809
16	151	62.4	426	21	AAV82207
17	151	62.4	426	21	AAV52681
18	151	62.4	426	21	AAV52172
19	151	62.4	426	21	AAV52976
20	140	57.9	426	15	AAV48068
21	137	56.6	424	15	AAV60165
22	137	56.6	424	15	AAV47198
23	136	56.2	424	15	AAV55202
24	136	56.2	424	20	AAV42475
25	136	56.2	424	20	AAV81812
26	136	56.2	424	21	AAV82210
27	136	56.2	424	21	AAV52684
28	136	56.2	424	21	AAV52175
29	136	56.2	424	21	AAV52979
30	126	52.1	424	11	AAV07058
31	119	49.2	258	14	AAV41004
32	119	49.2	258	15	AAV96951
33	119	49.2	420	15	AAV96950
34	119	49.2	421	15	AAV5196
35	119	49.2	421	20	AAV42469
36	119	49.2	421	20	AAV81806
37	119	49.2	421	21	AAV82204
38	119	49.2	421	21	AAV52678
39	119	49.2	421	21	AAV52169
40	119	49.2	421	21	AAV52973
41	109	45.0	421	15	AAV55205
42	109	45.0	421	20	AAV42478
43	109	45.0	421	20	AAV81815
44	109	45.0	421	21	AAV82213
45	109	45.0	421	21	AAV52687

#### ALIGNMENTS

RESULT 1  
ID AAV30214 standard; peptide: 41 AA.  
XX AAV30214;  
XX AAV30214;  
DT 01-NOV-1999 (first entry)  
XX  
DE Amino acids 308 to 348 of human zona pellucida protein 3 (hzp3).  
XX Human; zona pellucida protein 3; hzp3; glycoprotein; human spermatozoa;  
KW acrosome reaction; fertility; sperm motility; fertilization.  
XX  
OS Homo sapiens.  
XX  
PN M09942581-A1.  
XX  
PD 26-AUG-1999.  
XX  
PF 19-FEB-1999; 99WO-US03273.  
XX  
PR 19-FEB-1998; 98US-0075079.  
XX  
XX (EVIR-) EASTERN VIRGINIA MEDICAL SCHOOL.  
XX Dong K, Gibbons WE, Oehninger S;  
XX WPI: 1999-527476/44.  
XX  
XX New recombinant human zona pellucida protein 3 - used to develop  
XX products for diagnostic and therapeutic use related to human  
XX fertility  
XX  
PS Claim 7; Fig 1; 35pp: English.





```

XX 07-JUN-1995; 95US-0484993.
PF 09-NOV-1993; 93US-0149223.
XX 09-NOV-1992; 92US-0973341.
PR 29-JAN-1993; 93US-0012990.
XX 07-JUN-1995; 95US-0484993.
PA (ZONA-) ZONAGEN INC.
XX Harris JD;
PI WPI; 1999-023447/02.
DR N-PSDB; AAV64820.
XX Isolated zona pellucida DNA from different mammals - used to develop
PT products which can be used for vaccination to induce transient
PT infertility or permanent sterility in female mammals
XX Example 12; Column 145-148; 84pp; English.
XX This sequence represents a Macaca cynomolgus ZPC protein isolated from
CC zona pellucida. This protein can be used in a method for specifically
CC inducing transient infertility or permanent sterility in a host
CC animal by selective vaccination with specific zona pellucida proteins
CC or immunoontraceptively active fragments.
XX Sequence 223 AA;
SQ
Query Match 93.4%; Score 226; DB 20; Length 223;
Best Local Similarity 90.2%; Pred. No. 4e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMFPVQGPADICCCCKNGDCGTPSHSRROPHVMSQWSRSYS 41
107 swfvegpadiqccksgdgtfshrrghvsgwsrsas 147
RESULT 7
AA82218
ID AAY82218 standard; Protein; 223 AA.
XX AC AAY82218;
XX 12-JUN-2000 (first entry)
XX DE Monkey zona pellucida ZPC SEQ ID NO:49.
XX KW Zona pellucida; ZPA; ZPB; ZPC; infertility; permanent sterility;
XX vaccine; immunoontraceptive; contraceptive; immunostimulatory.
XX OS Primate.
XX US6027727-A.
XX 22-FEB-2000.
XX 09-NOV-1993; 93US-0149223.
XX 09-NOV-1992; 92US-0973341.
XX 29-JAN-1993; 93US-0012990.
XX PA (ZONA-) ZONAGEN INC.
XX PI Podolski JS, Hsu KT, Harris JD;
XX WPI; 2000-269144/23.
XX DR N-PSDB; AA295677.
XX Inducing reproducible transient or permanent infertility in a mammal
PT comprises administration of homologous and/or heterologous mammalian
PT species zona pellucida proteins

```

```

XX Example 12; Column 147-148; 85pp; English.
PS The present invention describes a method for inducing reproducible
XX transient infertility in a female mammal, including humans, by selective
CC administration of homologous and/or heterologous mammalian species
CC zona pellucida (ZP) proteins or immunoontraceptively active fragments.
CC Also described are: (1) a method for inducing transient infertility in a
CC mammal by administering directly an antibody directed to a ZPB or an
CC immunologically active fragment selected from feline, bovine,
CC cynomolgus monkey or human ZPB polypeptides; (2) an isolated, purified
CC recombinant feline, bovine, cynomolgus monkey or human ZPB polypeptide
CC or immunologically active fragment; and (3) a fusion protein comprising
CC a ZPB polypeptide which is conjugated with a compound selected from
CC keyhole limpet haemocyanin, muramyl dipeptide, histidine-tag, beta-gal,
CC or palmitic acid where the fusion protein remains effective to stimulate
CC production of antibodies that recognise a ZPB polypeptide. The method is
CC useful for providing transient or permanent infertility or sterility in
CC humans and animals. The present sequence represents monkey ZPC, which is
CC used in the exemplification of the present invention.
XX Sequence 223 AA;
SQ
Query Match 93.4%; Score 226; DB 21; Length 223;
Best Local Similarity 90.2%; Pred. No. 4e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMFPVQGPADICCCCKNGDCGTPSHSRROPHVMSQWSRSYS 41
107 swfvegpadiqccksgdgtfshrrghvsgwsrsas 147
DB 14-MAR-2000 (first entry)
RESULT 8
AA52692
ID AAY52692 standard; Protein; 223 AA.
XX AC AAY52692;
XX 14-MAR-2000 (first entry)
XX DE Cynomolgus monkey oocyte zona pellucida protein ZPC.
XX KW Zona pellucida; ZPA; ZPB; ZPC; purified; mammalian; glycoprotein;
XX target; immunoontraception; vaccine; antibody; transient; infertility;
XX controllable; predictable; permanent; sterility; side effect.
XX OS Macaca cynomolgus.
XX US6001599-A.
XX 14-DEC-1999.
XX 02-JUN-1995; 95US-0458731.
XX 09-NOV-1993; 93US-0149223.
XX 09-NOV-1992; 92US-0973341.
XX 29-JAN-1993; 93US-0012990.
XX PA (ZONA-) ZONAGEN INC.
XX PI Podolski JS, Hsu KT, Harris JD;
XX WPI; 2000-061880/05.
XX Isolated DNA encoding mammalian zona pellucida proteins useful for
PT inducing transient sterility -
XX Example 12; Columns 145-146; 84pp; English.
XX This sequence represents a mammalian zona pellucida protein, ZPC.
CC The zona pellucida is a complex matrix surrounding the mammalian
CC oocyte, formed of glycoproteins secreted by ovarian cells. The

```

invention relates to zona pellucida proteins ZPA, ZPB and ZPC, which are useful as targets for immunocontraceptive vaccines. Polynucleotides encoding these proteins can be used for recombinant protein production, and as probes in hybridisation methods for the isolation of polynucleotides encoding zona pellucida proteins from other mammalian species. Administration of zona pellucida proteins to a host animal results in a specific immunocontraceptive effect. Administration of purified ZPA and/or ZPB induces transient infertility in females. Fertility can be maintained in an on or off condition in a controllable/predictable fashion. Administration of ZPC induces permanent sterility. Use of a purified zona pellucida protein rather than a complex mixture of zona pellucida proteins reduces the potential for unwanted side-effects which may be harmful.

Sequence 223 AA;

Query Match 93.4%; Score 226; DB 21; Length 223;  
Best Local Similarity 90.2%; Pred. No. 4e-20; 1; Indels 0; Gaps 0;  
Matches 37; Conservative 3; Mismatches 1;

1 SWPVGADICCCNKGDCGTPSHRQPHVMSQMSRSVS 41  
||||:|||||||:|||||||:|||||||:|||||||  
Db 107 swfpgvpadicgccksgdcgtpshrrqphvsgwsrsas 147

RESULT 9

AA52183  
ID AAY52183 standard; Protein; 223 AA.

AC AAY52183;

D7 09-FEB-2000 (first entry)

X2 Monkey zona pellucida protein C (ZPC) amino acid sequence.

X1 Zona pellucida protein; ZPA; ZPB; ZPC; oocyte; ovarian cell; antibody;

KW fusion protein; vaccine; contraceptive; fertilization; fertility;

KM immunocontraceptive.

X3 Macaca cynomolgus.

PN US5989550-A.

PD 23-NOV-1999.

PF 07-JUN-1995; 95US-0480150.

PR 09-NOV-1993; 93US-0149223.

PR 09-NOV-1992; 92US-0973341.

PR 29-JAN-1993; 93US-0012990.

PA (ZONA-) ZONAGEN INC.

PI Podolski JS, Hsu KT, Harris JD;

DR WPI: 2000-022762/02.

DR N-PSDB; AA237824.

X1 Isolated and purified mammalian zona pellucida proteins useful in

PT methods of immunocontraception -

X3 Example 12; Column 145-146; 84pp; English.

CC Sequences AAY52181-Y52183 are monkey zona pellucida proteins ZPA, ZPB,

CC and ZPC. The invention relates to isolated and purified zona pellucida

CC proteins from mammals, specifically canine, feline and bovine ZPA or

CC their biologically active fragments also porcine and cynomolgus monkey

CC ZPA+. The zona pellucida is a complex matrix surrounding the mammalian

CC oocyte, formed of glycoproteins secreted by ovarian cells. The

CC previously mentioned ZPA proteins can be used to produce a fusion

CC protein to stimulate production of antibodies in a mammal that recognize

CC a ZPA polypeptide of the mammal. The ZPA polypeptides are useful as

CC vaccines for selectively effecting transient infertility in mammals. ZP

CC has an important role in maintaining the oocyte and in oocyte-sperm

CC interactions and so is a target for a contraceptive agent which

CC interferes with the fertilization process. Providing a specific

CC polypeptide reduces the need for a complex mixture of many ZP proteins

CC which may create unwanted and harmful side effects. The duration of

CC transient fertility is controllable and can be maintained in a

CC controllable and/or predictable fashion.

Sequence 223 AA;

Query Match 93.4%; Score 226; DB 21; Length 223;  
Best Local Similarity 90.2%; Pred. No. 4e-20; 1; Indels 0; Gaps 0;  
Matches 37; Conservative 3; Mismatches 1;

1 SWPVGADICCCNKGDCGTPSHRQPHVMSQMSRSVS 41  
||||:|||||||:|||||||:|||||||:|||||||  
Db 107 swfpgvpadicgccksgdcgtpshrrqphvsgwsrsas 147

RESULT 10

AA52987  
ID AAY52987 standard; Protein; 223 AA.

AC AAY52987;

D7 14-FEB-2000 (first entry)

X2 Cynomolgus monkey zona pellucida ZPC protein sequence.

X1 Zona pellucida; ZPA; ZPB; ZPC; infertility; sterility; vaccine;

KW immunocontraceptive; contraception; immunisation.

OS Mammalia.

PN US5976545-A.

PD 02-NOV-1999.

PF 07-JUN-1995; 95US-0484158.

PR 09-NOV-1993; 93US-0149223.

PR 09-NOV-1992; 92US-0973341.

PR 29-JAN-1993; 93US-0012990.

PA (ZONA-) ZONAGEN INC.

PI Podolski JS, Hsu KT, Harris JD;

DR WPI: 2000-037080/03.

DR N-PSDB; AA233274.

X1 New recombinant mammalian zona pellucida C proteins, for induction of

PT female sterility -

PS Claim 14; Column 145-146; 87pp; English.

CC The present invention describes recombinant zona pellucida (ZP)

CC proteins, specifically ZPC proteins. Also described are fusion proteins

CC of ZPC with any of keyhole limpet haemocyanin, muramyl dipeptide,

CC histidine tag, beta-galactosidase or palmitic acid, capable of

CC stimulating mammalian antibodies that recognise mammalian ZPC.

CC Recombinant ZPC proteins are used to immunise animals to induce transient

CC infertility or permanent female sterility, in humans or other mammals.

CC AA233243 to AA233254, AA233270 to AA233274 and AA233285 encode mammalian

CC ZP proteins given in AAY52971 to AAY52988 from the present invention.

CC AA233255 to AA233269 and AA233275 to AA233284 represent oligonucleotides

CC used in the exemplification of the present invention.

Sequence 223 AA;

Query Match 93.4%; Score 226; DB 21; Length 223;

Best Local Similarity 90.2%; Pred. No. 4e-20;

Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWPVGADICCCCNKGDGTPSHSRQPHVMSQMSRSVS 41

DB 107 swfvegpadiqcckskgdcgtpshsrqphvmsqmsrsas 147

#### RESULT 11

AA022239 AAR22239 standard; Protein; 372 AA.

AC AAR22239;

DT 10-JUL-1992 (first entry)

DE Human ZP3.

KW Zona pellucida; antigenic; epitopes; fertilisation; gametes; immunocontraception.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Peptide 341..360 /note="antigenic epitope"

XX MO9203548-A.

XX PD 05-MAR-1992.

XX PF 13-AUG-1991; 91WO-EP01538.

XX PR 27-AUG-1990; 90EP-0202287.

XX PA (ALKU ) AKZO NV.

XX PI Van Duin M;

XX DR WPI: 1992-096892/12.

XX DR N-PSDB; AA022997.

XX PT Human zona pellucida protein ZP3 and fragments - useful as

XX PS contraceptive vaccines or diagnostic medical kits

XX PS Claim 3; Fig 2; 32pp; English.

CC The protein sequence was deduced from the nucleotide sequence of

CC zona pellucida ZP3 obtd. by screening a human genomic EMBL3 library

CC with a labelled ZP3 probe. The human gene was found to comprise

CC 8 exons spread over 20 kb genomic DNA. cDNA clones were isolated

CC and expressed in CHO cells. Elucidation of the ZP3 sequence will

CC allow prodn. of large amts. of the protein by recombinant techniques.

CC The protein and derived polypeptides may be used to prep. an immuno-

CC contraction by either active or passive immunisation. The

CC polypeptide and antibodies (and anti-idiotypic antibodies) directed

CC to it may be useful in diagnostic kits.

XX SO Sequence 372 AA;

Query Match 93.0%; Score 225; DB 13; Length 372;

Best Local Similarity 92.5%; Pred. No. 8.7e-20;

Matches 37; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 WFPVQGPADICCCCNKGDGTPSHSRQPHVMSQMSRSVS 41

DB 309 wfpvegpadiqcckskgdcgtpshsrqphvmsqmsrsas 348

#### RESULT 12

AA053498

ID AAR53498 standard; Protein; 424 AA.

XX AAR53498;

XX AC 02-NOV-1994 (first entry)

XX DE Marmoset ZP3.

KW Zona pellucida; glycoprotein 3; sperm receptor; immunocontraceptive;

KW diagnosis; tumour; autoimmune disease.

OS Callithrix jacchus.

XX PN WO9410304-A.

XX PD 11-MAY-1994.

XX PF 29-OCT-1993; 93WO-EP03014.

XX PR 02-NOV-1992; 92EP-0310026.

XX PA (AKZU ) AKZO NOBEL NV.

XX PI Aitken RJ, Kootan T;

XX DR WPI: 1994-167461/20.

XX DR N-PSDB; AA063673.

XX PT New polypeptide(s) with marmoset ZP3 activity - are useful as

XX PT immuno-contraceptive vaccines and in diagnosis, also test system

XX PT for contraceptive development

XX PS Claim 1; Page 15; 27pp; English.

CC The marmoset ZP3 gene was isolated by PCR from total marmoset RNA

CC using PCR primers corresp. to regions of the human ZP3 gene. ZP3

CC (the sperm receptor) is useful as an immunocontraceptive and as a

CC diagnostic reagent e.g. to detect sperm function, to monitor the

CC effects of vaccination, for imaging follicles, to detect ZP3-related

CC autoimmune disorders or ZP3-expressing tumours or to screen for

CC autoantibodies against ZP3.

XX SO Sequence 424 AA;

Query Match 79.3%; Score 192; DB 15; Length 424;

Best Local Similarity 82.1%; Pred. No. 1e-15;

Matches 32; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SWPVGADICCCCNKGDGTPSHSRQPHVMSQMSRS 39

DB 308 swfvegpadiqcckskgdcgtpshsrqphvmsqmsrs 346

#### RESULT 13

AA055199 AAR55199 standard; Protein; 426 AA.

XX AC AAR55199;

XX DT 31-JAN-1995 (first entry)

XX DE Canine zona pellucida ZPC protein.

KW Dog; canine; zona pellucida; ZPC; immunocontraception.

XX OS Canis familiaris.

XX OS Key

XX FT Location/Qualifiers

XX FT Protein 1..426

XX FT /label= canine\_ZPC

XX PN WO9411019-A.

[illegible]





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 08:24:10 ; Search time 24.08 Seconds  
(without alignments)  
34.299 Million cell updates/sec

Title: US-09-252-828A-2  
Perfect score: 242  
Sequence: 1 SMFVPGADICQCCNKGDG.....TPSHRRQPHVMSQMSRVS 41

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues  
Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	235	97.1	424	1 US-08-453-472-7	Sequence 7, Appl
2	235	97.1	424	1 US-08-038-948-4	Sequence 4, Appl
3	235	97.1	424	1 US-08-453-952-7	Sequence 7, Appl
4	235	97.1	424	2 US-08-862-903-7	Sequence 7, Appl
5	235	97.1	424	2 US-08-484-158B-61	Sequence 61, Appl
6	226	93.4	223	2 US-08-484-93B-49	Sequence 49, Appl
7	226	93.4	223	2 US-08-484-158B-49	Sequence 49, Appl
8	226	93.4	223	2 US-08-484-596A-49	Sequence 49, Appl
9	226	93.4	223	2 US-08-480-150A-49	Sequence 49, Appl
10	226	93.4	223	3 US-08-458-731-49	Sequence 49, Appl
11	226	93.4	223	3 US-08-149-223A-49	Sequence 49, Appl
12	226	93.4	426	2 US-08-484-93B-12	Sequence 12, Appl
13	226	93.4	426	2 US-08-484-158B-12	Sequence 12, Appl
14	226	93.4	426	2 US-08-484-596A-12	Sequence 12, Appl
15	226	93.4	426	2 US-08-480-150A-12	Sequence 12, Appl
16	226	93.4	426	3 US-08-458-731-12	Sequence 12, Appl
17	226	93.4	426	3 US-08-149-223A-12	Sequence 12, Appl
18	226	93.4	426	2 US-08-484-93B-18	Sequence 18, Appl
19	226	93.4	426	2 US-08-484-158B-18	Sequence 18, Appl
20	226	93.4	426	2 US-08-484-596A-18	Sequence 18, Appl
21	226	93.4	426	2 US-08-480-150A-18	Sequence 18, Appl
22	226	93.4	426	3 US-08-458-731-18	Sequence 18, Appl
23	226	93.4	426	3 US-08-149-223A-18	Sequence 18, Appl
24	226	93.4	426	1 US-08-453-472-8	Sequence 8, Appl
25	226	93.4	426	1 US-08-038-948-2	Sequence 2, Appl
26	226	93.4	426	1 US-08-038-948-5	Sequence 5, Appl
27	226	93.4	426	1 US-08-453-952-8	Sequence 8, Appl

28	126	52.1	424	2 US-08-862-903-8	Sequence 8, Appl
29	119	49.2	421	2 US-08-484-93B-6	Sequence 6, Appl
30	119	49.2	421	2 US-08-484-158B-6	Sequence 6, Appl
31	119	49.2	421	2 US-08-484-596A-6	Sequence 6, Appl
32	119	49.2	421	2 US-08-480-150A-6	Sequence 6, Appl
33	119	49.2	421	3 US-08-458-731-6	Sequence 6, Appl
34	119	49.2	421	3 US-08-149-223A-6	Sequence 6, Appl
35	109	45.0	421	2 US-08-484-93B-24	Sequence 24, Appl
36	109	45.0	421	2 US-08-484-158B-24	Sequence 24, Appl
37	109	45.0	421	2 US-08-484-596A-24	Sequence 24, Appl
38	109	45.0	421	2 US-08-480-150A-24	Sequence 24, Appl
39	109	45.0	421	3 US-08-458-731-24	Sequence 24, Appl
40	109	45.0	421	3 US-08-149-223A-24	Sequence 24, Appl
41	100	41.3	415	2 US-08-484-93B-8	Sequence 8, Appl
42	100	41.3	415	2 US-08-484-158B-8	Sequence 8, Appl
43	100	41.3	415	2 US-08-484-596A-8	Sequence 8, Appl
44	100	41.3	415	2 US-08-480-150A-8	Sequence 8, Appl
45	100	41.3	415	3 US-08-458-731-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
Sequence 7, Application US/08453472  
Patent No. 5626846  
GENERAL INFORMATION:  
APPLICANT: DEAN, JURRIEN  
TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
TITLE OF INVENTION: BASED ON ALLOIMUNIZATION WITH ZONA PELLUCIDA  
TITLE OF INVENTION: POLYPEPTIDES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,472  
FILING DATE: 30-May-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/038,948  
FILING DATE: 26-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/930,462  
FILING DATE: 20-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/364,379  
FILING DATE: 12-JUN-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4032 US3  
TELEPHONE: (212) 751-6849  
TELEFAX: (212) 751-6849  
TELEEX: 421792  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein

```

ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: ZP3
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: human ZP3 protein
US-08-453-472-7

Query Match 97.1%; Score 235; DB 1; Length 424;
Best Local Similarity 95.1%; Pred. No. 6.4e-21;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWPVGADICCCCKNGDCGPHSRROPHVMSQMSRSYS 41
Db 308 SWPVGADICCCCKNGDCGPHSRROPHVMSQMSRSAS 348

RESULT 2
US-08-038-948-4
Sequence 4, Application US/08038948
Patent No. 564187
GENERAL INFORMATION:
APPLICANT: DEAN, JURRIEN
TITLE OF INVENTION: CONTRACEPTIVE VACCINE BASED ON
TITLE OF INVENTION: ALLOIMMUNIZATION WITH ZONA PELLUCIDA POLYPEPTIDES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/038,948
FILING DATE: 26-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: 99152/E-266-88/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

```

```

US-08-038-948-4
Query Match          97.1%, Score 235; DB 1; Length 424;
Best Local Similarity 95.1%, Pred. NO. 6.4e-21;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy      1  SWFPGPADICCCCNKGDCTPSHSRRORPHYMSQMSRSVS 41
        |||||
Db      308 SWFPGPADICCCCNKGDCTPSHSRRORPHYMSQMSRSVS 348

RESULT      3
US-08-453-952-7
; Sequence 7, Application US/08453952
; Patent No. 5672488
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMUNIZATION WITH ZONA PELLUCIDA
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,952
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 US4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPOLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:

```

```

?      LENGTH: 424 amino acids
?      TYPE: amino acid
?      TOPOLOGY: linear
?      MOLECULE TYPE: protein
?      DESCRIPTION:
?      DESCRIPTION: "deduced amino acid sequence of human ZPC

```

500

CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,596A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/149,223  
FILING DATE: 11-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-596A-49

Query Match 93.4%; Score 226; DB 2; Length 223;  
Best Local Similarity 90.2%; Pred. No. 4.1e-20;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWFVPGPADICCCCKNGDCGTPSHSRQPHVMSQMSRSVS 41  
|||||  
Db 107 SWFVEGPADICCCCKNGDCGTPSHSRQPHVMSQMSRSVS 147

RESULT 9  
US-08-480-150A-49  
Sequence 49, Application US/08480150A  
Patent No. 5989550  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,150A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,223

FILING DATE: 09-NOV-1993  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-480-150A-49

Query Match 93.4%; Score 226; DB 2; Length 223;  
Best Local Similarity 90.2%; Pred. No. 4.1e-20;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWFVPGPADICCCCKNGDCGTPSHSRQPHVMSQMSRSVS 41  
|||||  
Db 107 SWFVEGPADICCCCKNGDCGTPSHSRQPHVMSQMSRSVS 147

RESULT 10  
US-08-458-731-49  
Sequence 49, Application US/08458731  
Patent No. 6001599  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,731  
FILING DATE: 09-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 49:

```

; SEQUENCE CHARACTERISTICS:
;     LENGTH: 223 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-458-731-49

```

Query Match	93.48;	Score 226;	DB 3;	Length 223
Best Local Similarity	90.28;	Pred. No. 4.1e-20;		
Matches 37; Conservative	3;	Mismatches 1;	Indels	

```

QY      1  SWEPRVGGPADICCCCKNGDCGTPSHSRROQHVMSSQWSRSV  41
        |||||:|||||||:|||||||:|||||||:|||||||:
Db     107  SWEPRVEGPADICCCCKNGDCGTPSHSRROQHVMSSQWSRSAS  147

```

US-08-149-223A-49  
SEQUENCE 49, Application US/08149223A  
Patent No. 6027727

GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immuncontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/149,223A  
FILING DATE: 09-NOV-1993  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3836

INFORMATION FOR SEQ ID NO.: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-149-223A-49

	Query Match	93.48;	Score 226;	DB 3;	Length 223;
	Best Local Similarity	90.28;	Pred. No. 4.1e-20;		
	Matches 37;	Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0;
1	SWPVGSGPADIQCQCNNGDCGTPSHSRPQHYVMSQMSRSVS	41			
	:     :				
107	SWPVGSGPADIQCQCSKGDGCTPSHSRPPHYVMSQMSRSAS	147			

RESULT 12  
US-08-484-993B-12  
; Sequence 12, Application US/08484993B  
; Patent No. 5837497  
GENERAL INFORMATION

Query Match	62.48;	Score 151;	DB 2;	Length 426;
Best Local Similarity	60.08;	Pred. No. 7.5e-11;		
Matches 24;	Conservative	5;	Mismatches 11;	Indels 0;
				Gaps 0;

QY 2 WEPVGGPADICGCCCKMGDCGTPSHSRQPHYMSQWMSVS 41  
|:::| ::::|::| | | | | : |::|  
Dbd 307 WYPVGSAIDICRCCKNGSGCLPGRSRLSLHLEGRKRSVS 346

RESULT 13  
US-08-484-158B-12  
Sequence 12, Application US/08484158B  
Patent No. 5976545  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Pharmaceutical Compositions for  
TITLE OF INVENTION: Immunocotraception  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

```
Q?      2 WFPYQGADIDCCCKGDCGTPSHSRRCPHYMSOMSRSVS 41  
        I::I: I::II:I::II::II::I: I::I: I::I:  
DB      307, WYPEGSADICRCCKNGSGLPGRSRRLSHLERGWKRSVS 346
```

---

```
RESULT   14  
US-08-484-596A-12  
; Sequence 12, Application US/08484596A  
; Patent No. 5981228  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum  
STREET: 6500 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:
```

Query Match	62.4%	Score 151	DB 2:	Length 426;
Best Local Similarity	60.0%	Pred. No.	7.5e-11;	
Matches 24;	Conservative	5;	Mismatches 11;	Indels 0;
				Gaps 0;

RESULT 15

Sequence 12, Application US/08480150A  
Patent No. 5989550  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,150A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,223  
FILING DATE: 09-NOV-1993  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David M.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

04 protein - protein search, using sw model

Run on: July 12, 2001, 08:24:10 ; Search time 41.53 Seconds  
(without alignments)  
75.202 Million cell updates/sec

Title: US-09-252-828a-2

Perfect score: 242

Sequence: 1 SMFPVQGPADICQCCNKGDG.....TSHSRQPHVMSQMSRSVS 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 segs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	235	97.1	424	1	A36000	sperm-binding glyco
2	217	89.7	210	2	A56844	POM-zp3 protein -
3	151	62.4	426	2	S70396	zona pellucida gly
4	143	59.1	422	1	A60503	sperm-binding glyco
5	136	56.2	424	2	S70399	zona pellucida gly
6	126	52.1	424	1	A30334	sperm-binding glyco
7	119	49.2	421	1	S70433	zona pellucida gly
8	109	45.0	421	1	S70402	zona pellucida gly
9	100	41.3	415	2	S70401	zona pellucida gly
10	98	40.5	44	2	B44365	sperm receptor lig
11	60.5	25.0	99	2	B82720	hypothetical prote
12	59	24.4	58	2	S35373	zona pellucida gly
13	59	24.4	999	1	IJH063	desmoglein 3 precu
14	58.5	24.2	357	2	S09269	ig alpha chain C r
15	58	24.0	66	2	S78705	protein YBR056w-a
16	56.5	23.3	819	2	G81698	leucyl-tRNA synthet
17	56	23.1	647	2	JE0337	frizzled-1 protein
18	55.5	22.9	436	2	JN0591	serotonin receptor
19	55	22.7	641	2	A45054	probable intercell
20	55	22.7	858	2	T08881	prominin - mouse
21	55	22.7	865	2	T09050	AC133 antigen - hu
22	54.5	22.5	424	2	S52847	egg membrane prote
23	54.5	22.5	735	2	I48101	ADAM 6 protein pre
24	54	22.3	202	2	A33176	P21 protein - soyb
25	54	22.3	477	2	A47236	zinc-finger protei
26	54	22.3	494	2	A42170	zinc finger protei
27	54	22.3	497	2	JC5076	myc-associated zin
28	53.5	22.1	351	2	C72508	hypothetical prote
29	53.5	22.1	379	2	H96696	protein FIN21.16 l

30	53.5	22.1	738	2	T00343	hypothetical prote
31	53	21.9	197	2	I46413	keratin KAP5.5 - s
32	53	21.9	563	2	T47520	hypothetical prote
33	52.5	21.7	447	2	S39505	translation elonga
34	52.5	21.7	447	2	S17434	translation elonga
35	52.5	21.7	447	2	JS0719	translation elonga
36	52.5	21.7	447	2	JC1454	translation elonga
37	52.5	21.7	447	2	S66339	translation elonga
38	52.5	21.7	448	2	S10507	translation elonga
39	52.5	21.7	449	2	S08534	translation elonga
40	52.5	21.7	449	2	S06724	translation elonga
41	52.5	21.7	449	2	S21989	translation elonga
42	52.5	21.7	967	2	F86214	protein TGD22.2 li
43	52	21.5	127	2	T08097	osmotin - sweet or
44	51.5	21.3	339	2	F86614	phenylalanyl tRNA
45	51.5	21.3	339	2	B72011	phenylalanine--trn

#### ALIGNMENTS

RESULT 1  
A36000  
sperm-binding glycoprotein ZP3 precursor - human  
N:Alternate names: sperm receptor ZP3; zona pellucida glycoprotein ZP3  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 10-Sep-1999  
C:Accession: A36000; A44365  
R:Chamberlin, M.E.; Dean, J.  
Proc. Natl. Acad. Sci. U.S.A. 87, 6014-6018, 1990  
A:Title: Human homolog of the mouse sperm receptor.  
A:Reference number: A36000; MUID:90349545  
A:Accession: A36000  
A:Molecule type: mRNA; DNA  
A:Residues: 1-424 <CH>  
A:Cross-references: GB:M60504; GB:M35109; NID:9340491; PIDN:AAA61336.1; PID:9340492  
R:van Duin, M.; Polman, J.E.; Verkoelen, C.C.; Bunschoten, H.; Meyerink, J.H.; Olijve  
Genomics 14, 1064-1070, 1992  
A:Title: Cloning and characterization of the human sperm receptor ligand ZP3: evidenc  
A:Reference number: A44365; MUID:93122771  
A:Accession: A44365  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 329-370, 'S', 372-424 <VAN>  
A:Experimental source: ovary  
A:Note: sequence inconsistent with the nucleotide translation  
A:Note: sequence extracted from NCBI backbone (NCBI:122391, NCBI:122392)  
C:Comment: This sulfated glycoprotein in the zona pellucida of the oocyte is a recept  
C:Genetics:  
A:Gene: GDB:2P3A  
A:Cross-references: GDB:128007; OMIM:182889  
A:Map position: 7pter-7qter  
C:Keywords: glycoprotein; oocyte; receptor; sulfolipin; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-424/Product: sperm-binding glycoprotein ZP3 #status predicted <MAT>  
F:45-301/Domain: ZP domain homology <ZPH>

Query Match 97.1% Score 235; DB 1; Length 424;  
Best Local Similarity 95.1% Pred. No. 1.2e-20;  
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SMFPVQGPADICQCCNKGDGCTPSHSRQPHVMSQMSRSVS 41  
|||||  
Db 308 SMFPVQGPADICQCCNKGDGCTPSHSRQPHVMSQMSRSVS 348

RESULT 2  
A56844  
POM-zp3 protein - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Oct-1995 #sequence-revision 19-Oct-1995 #text-change 05-Nov-1999



A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-152, 'E', 154-252, 'E', 254-424 <ROS>

A:Experimental source: ovary

A:Note: sequence extracted from NCBI backbone (NCBIP:119137)

C:Comment: This sulfated glycoprotein in the zona pellucida of the oocyte is a receptor

C:Genetics:

A:Introns: 103/3; 143/2; 178/1; 239/2; 278/3; 309/2; 355/1

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

C:Keywords: glycoprotein; oocyte; sulfoprotein; transmembrane protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-424/Product: sperm-binding glycoprotein ZP3 #status predicted <MAT>

F:45-302/Domain: ZP domain homology <ZPH>

F:146,304,327,330/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:273/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match

Best Local Similarity 52.1%; Score 126; DB 1; Length 424;

Matches 22; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 SWPVOGPADICCCCKNGDCGTPSHSRORPHVMSQMSRSVS 41

DB 309 SWPVEGSDADICCCSHGNCSSSSQFOIHGPRQMSKLV 349

RESULT 7

zona pellucida glycoprotein C - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

A:Title: Cloning and characterization of zona pellucida genes

A:Reference number: S70396; MUID:95143578

A:Accession: S70433

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-421 <HAR>

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

F:44-300/Domain: ZP domain homology <ZPH>

Query Match

Best Local Similarity 49.2%; Score 119; DB 1; Length 421;

Matches 19; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 WFPVOGPADICCCCKNGDCGTPSHSR 28

DB 308 WSPVEGPATVRCCHGCGCTPSLSRK 334

RESULT 8

zona pellucida glycoprotein C - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

A:Title: Cloning and characterization of zona pellucida genes

A:Reference number: S70396; MUID:95143578

A:Accession: S70402

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-421 <HAR>

A:Cross-references: EMBL:U05775; NID:g458266; PIDN:AAA74385.1; PID:g458267

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

F:44-300/Domain: ZP domain homology <ZPH>

Query Match

45.0%; Score 109; DB 1; Length 421;

Best Local Similarity 56.7%; Pred. No. 1.1e-05;

Matches 17; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 2 WFPVOGPADICCCCKNGDCGTPSHSRORPH 31

DB 308 WSPVEGPTDICRCCKSGRCGSGRSMRLSH 337

RESULT 9

zona pellucida glycoprotein C - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 20-Aug-1999

A:Accession: S70401

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-415 <HAR>

A:Cross-references: EMBL:U05782; NID:g458280; PIDN:AAA74392.1; PID:g458281

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

F:41-295/Domain: ZP domain homology <ZPH>

Query Match

Best Local Similarity 41.3%; Score 100; DB 2; Length 415;

Matches 20; Conservative 5; Mismatches 10; Indels 12; Gaps 2;

OY 1 SWPVOGPADICCCCKNGDC-----GTP-----SHSRORPHVMSQ 35

DB 302 SWAPVEGSADICCCCKNGDCDLIAGSPNQNHRARSLSRRHVTVE 348

RESULT 10

B44365

sperm receptor ligand homolog ZP3-372 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 12-May-1995

A:Accession: B44365

A:Title: Cloning and characterization of the human sperm receptor ligand ZP3: evidence

A:Reference number: A44365; MUID:93122771

A:Accession: B44365

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-44 <VAN>

A:Experimental source: ovary

A:Note: sequence extracted from NCBI backbone (NCBIN:123950, NCBIP:122393)

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

Query Match

Best Local Similarity 40.5%; Score 98; DB 2; Length 44;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 22 TPSSRRORPHVMSQMSRSVS 41

DB 1 TPSSRRORPHVMSQMSSTVS 20

RESULT 11

B82720

hypothetical protein xrl139 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

A:Accession: B82720

A:Title: The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.



A: Reference number: S45906  
 A: Accession: S78705  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-66 <AlJ>  
 A: Cross-references: EMBL:Z35925; MIPS:YBR056w-a  
 A: Experimental source: strain S288C  
 C: Genetics:  
 A: Map position: 2R

Query Match 24.0%; Score 58; DB 2; Length 66;  
 Best Local Similarity 57.9%; Pred. No. 3.1;  
 Matches 11; Conservative 0; Mismatches 4; Indels 4; Gaps 1;  
 Q7 6 QGPA---DQCQCCKRGDC 20  
 ||| | | | | | |  
 D3 24 QGPPPPRNDCCCCCNCGDC 42

Search completed: July 12, 2001, 08:26:25  
 Job time: 135 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

CM protein - protein search, using sw model

Run on: July 12, 2001, 08:24:10 ; Search time 10.02 Seconds

(without alignments)  
140.167 Million cell updates/sec

Title: US-09-252-828A-2

Sequence: 1 SMFVQGRADICCCCKGDC.....TFSHRQPRHVSQMSKSVS 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_39:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	97.1	424	1	2P3A_HUMAN
2	229	94.6	372	1	2P3B_HUMAN
3	226	93.4	424	1	2P3A_MACRA
4	192	79.3	424	1	2P3A_CALSQ
5	151	62.4	426	1	2P3_CANFA
6	143	59.1	422	1	2P3_MESAU
7	136	56.2	424	1	2P3_FELCA
8	126	52.1	424	1	2P3_MOUSE
9	119	49.2	421	1	2P3_PIG
10	109	45.0	421	1	2P3_BOVIN
11	100	41.3	415	1	2P3_RABIT
12	59	24.4	999	1	DSG3_HUMAN
13	55.5	22.9	436	1	5H6_RAT
14	55	22.7	865	1	PMH1_HUMAN
15	55	22.7	867	1	PROM_MOUSE
16	54	22.3	202	1	P21_SOYBN
17	54	22.3	477	1	MAZ_HUMAN
18	54	22.1	477	1	MAZ_MOUSE
19	53.5	22.3	338	1	DEFB_HUMAN
20	52.5	21.7	447	1	EPH1_HORVU
21	52.5	21.7	447	1	EPH12_DAUCA
22	52.5	21.7	447	1	EPH12_HORVU
23	52.5	21.7	447	1	EPH12_HORVU
24	52.5	21.7	447	1	EPH12_HORVU
25	52.5	21.7	447	1	EPH12_HORVU
26	52.5	21.7	447	1	EPH12_HORVU
27	52.5	21.7	447	1	EPH12_HORVU
28	52.5	21.7	447	1	EPH12_HORVU
29	52.5	21.7	447	1	EPH12_HORVU
30	52.5	21.7	448	1	EPH12_HORVU
31	52.5	21.7	448	1	EPH12_HORVU
32	52.5	21.7	449	1	EPH12_HORVU
33	52.5	21.7	449	1	EPH12_HORVU

34	51.5	21.3	339	1	SYFA_CHLPN	09%6r6 chlamydia p
35	51	21.1	65	1	TRW6_NAJNA	P29180 naja naja (
36	51	21.1	281	1	TRYZ_DROBR	P54630 drosophila
37	51	21.1	1132	1	DNEI_HSV6U	P52338 human herpes
38	51	21.1	1132	1	DNEI_HSV6Z	P52538 human herpes
39	50.5	20.9	74	1	NXLD_ACAAN	P34073 acanthophis
40	50.5	20.9	139	1	IGF_MYXGL	P22618 myxine glut
41	50.5	20.9	622	1	ACHA_CHICK	P09482 gallus gall
42	50	20.7	174	1	SODC_CAUCR	P20379 caulobacter
43	50	20.7	182	1	KRUC_SHEEP	P26372 ovis aries
44	50	20.7	269	1	ILIB_MOUSE	P10749 mus musculu
45	50	20.7	643	1	RHOP_MOUSE	O61085 mus musculu

## ALIGNMENTS

RESULT 1	STANDARD:	PRT: 424 AA.
ID 2P3A_HUMAN		
AC P21754;		
DT 01-MAY-1991 (Rel. 18, Created)		
DT 01-MAY-1991 (Rel. 18, Last sequence update)		
DT 30-MAY-2000 (Rel. 39, Last annotation update)		
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A PRECURSOR (ZONA PELLUCIDA		
DE GLYCOPROTEIN ZP3A) (ZONA PELLUCIDA PROTEIN C) (SPERM RECEPTOR) (ZP3).		
GN ZP3A OR ZP3.		
OS Homo sapiens (Human).		
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX NCBI_TaxID=9606;		
RN [1]		
RP SEQUENCE FROM N.A.		
RX MEDLINE=90349545; PubMed=2385582;		
RA Chamberlin M.E., Dean J.;		
RT "Human homolog of the mouse sperm receptor.";		
RL Proc. Natl. Acad. Sci. U.S.A. 87:6014-6018(1990).		
RN [2]		
RP SEQUENCE OF 329-424 FROM N.A.		
RC TISSUE=Ovary;		
RX MEDLINE=93122771; PubMed=1478648;		
RA van Duin M., Polman J.E., Verkoelen C.C., Bunschoten H.,		
RA Meyerink J.H., Olijve W., Aitken R.J.;		
RT "Cloning and characterization of the human sperm receptor ligand ZP3:		
RT in the Caucasian and Japanese populations.";		
RL Genomics 14:1064-1070(1992).		
CC - FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR		
CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE		
CC SPECIES-SPECIFICITY OF THE INSEMINATION.		
CC - SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN		
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.		
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR		
CC MATRIX.		
CC - PFM: SUFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.		
CC - SIMILARITY: CONTAINS 1 ZP DOMAIN.		
CC -----		
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -		
CC the European Bioinformatics Institute. There are no restrictions on its		
CC use by non-profit institutions as long as its content is in no way		
CC modified and this statement is not removed. Usage by and for commercial		
CC entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC -----		
CC EMBL: M60504; AAA61336.1; -		
CC PIR: A36000; A36000.		
CC MIM: 182889; -		
CC InterPro: IPR001507; -		
CC Pfam: PF00100; zona_pellucida; 1.		
CC PRINTS: PR00023; ZPELUCIDA.		
CC PROSITE: PS00682; ZP_DOMAIN; 1.		
CC Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;		





FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 424 AA; 47040 MW; 3B41C4CFA3792331 CRC64;

Query Match  
 Best Local Similarity 90.2%; Score 226; DB 1; Length 424;  
 Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWPVGADICCCCKGDCGTPSHRQPHVMSQMSRSVS 41  
 DB 308 SWPVGADICCCCKGDCGTPSHRQPHVMSQMSRSVS 348

RESULT 4  
 ID ZP3A\_CALSO STANDARD; PRT; 424 AA.  
 AC P53786;  
 DT 01-FEB-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP3A) (ZONA PELLUCIDA PROTEIN C) (SPERM RECEPTOR) (ZP3).  
 GN ZP3A OR ZP3.  
 OS Callithrix sp. (Marmoset).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;  
 OC NCBI\_TaxID=9485;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Ovary;  
 RX MEDLINE=94363314; PubMed=8081814;  
 RA Thilak-Koother P., Van Duin M., Aitken R.J.;  
 RT "Cloning, sequencing and oocyte-specific expression of the marmoset sperm receptor protein, ZP3.";  
 RL Zygote 1:93-101(1993).  
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE SPECIES-SPECIFICITY OF THE INSEMINATION.  
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR MATRIX.  
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.  
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: S71825; AAB31866.1; -;  
 DR InterPro: IPR001507; -;  
 DR Pfam: PF00100; zona\_pellucida; 1.  
 DR PRINTS: PR00023; ZPELLUCIDA.  
 DR PROSITE: PS00682; ZP\_DOMAIN; 1.  
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane; Extracellular matrix; Multigene family;  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A.  
 FT DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 388 408 POTENTIAL.  
 FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 45 307 ZP.  
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 424 AA; 46809 MW; 1DACBD03026C2739 CRC64;

Query Match  
 Best Local Similarity 79.3%; Score 192; DB 1; Length 424;  
 Matches 32; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SWPVGADICCCCKGDCGTPSHRQPHVMSQMSRS 39  
 DB 308 SWPVGADICCCCKGDCGTPSHRQPHVMSQMSRS 346

RESULT 5  
 ID ZP3\_CANFA STANDARD; PRT; 426 AA.  
 AC P48831;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).  
 GN ZP3 OR ZPC.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OC NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Ovary;  
 RX MEDLINE=95143578; PubMed=7841460;  
 RA Harris J.D., Hübner D.W., Fontenot G.K., Hsu K.T., Virewicz E.C., Sacco A.G.;  
 RT "Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";  
 RL DNA Seq. 4:361-393(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Ovary;  
 RA Okazaki Y., Sugimoto M.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).  
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR MATRIX.  
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U05780; AAA74387.1; -;  
 DR EMBL: DA5070; BAA08098.1; -;  
 DR InterPro: IPR001507; -;  
 DR Pfam: PF00100; zona\_pellucida; 1.  
 DR PRINTS: PR00023; ZPELLUCIDA.  
 DR PROSITE: PS00682; ZP\_DOMAIN; 1.  
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane; Extracellular matrix; Multigene family;  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 426 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.  
 FT DOMAIN 23 385 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 386 406 POTENTIAL.  
 FT DOMAIN 407 426 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 43 305 ZP.

DR EMBL, 005778; AAA74390.1; -;  
DR EMBL, D45068; BAA08096.1; -;  
DR InterPro, IPR001507; -;  
DR Pfam, PF00100; zona pellucida. 1.  
DR PRINTS, PR00023; zPELUCIDA.  
DR PROSITE, PS00682; ZP\_DOMAIN; FALSE\_NEG.  
KW Glycoprotein; Signal; Sulfatase; Sperm;  
KW Extracellular matrix; Multigene family;  
KW Receptor; Transmembrane;

```

FE SIGNAL 1 22 POTENTIAL.
FE CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
FE DOMAIN 23 383 EXTRACELLULAR (POTENTIAL).
FE TRASNEM 384 404 POTENTIAL.
FE DOMAIN 405 424 CYTOPLASMIC (POTENTIAL).
FE DOMAIN 43 305 ZP.
FE CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FE CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FE CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FE CONFLICT 72 72 D -> W (IN REF. 2).
FE CONFLICT 264 264 D -> Y (IN REF. 2).
SQ SEQUENCE 424 AA: 46853 MW: CFC62F35F9AFC0D CRC64;

Query Match 56.2%; Score 136; DB 1; Length 424;
Best Local Similarity 55.0%; Pred. No. 1,8e-09;
Matches 22; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 2 WFPVGGPADICCCCKGCGTSPHSRRQPHVMSQMSRSVS 41
DB 307 WFPVGGPADICCCCKGCGTSPHSRRQPHVMSQMSRSVS 346

RESULT 8
ZP3_MOUSE STANDARD; PRT; 424 AA.
ID ZP3_MOUSE
AC P10761.
ET 01-JUL-1989 (Rel. 11, Created)
ET 01-NOV-1997 (Rel. 35, Last sequence update)
ET 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
CN ZP3 OR ZP-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88242926; PubMed=3378665;
RX Ringuette M.J., Chamberlin M.E., Baur A.W., Sobleski D.A., Dean J.;
RT "Molecular analysis of cDNA coding for ZP3, a sperm binding protein
RT of the mouse zona pellucida."
RL Dev. Biol. 127:287-295(1988).
RN [2]
RP REVISION TO 387.
RA Dean J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=CD-1; TISSUE=Liver;
RX MEDLINE=89240048; PubMed=2541416;
RA Kinloch R.A., Rollier R.J., Fimiani C.M., Wassarman D.A.;
RT "Primary structure of the mouse sperm receptor polypeptide determined
RT by genomic cloning."
RL Proc. Natl. Acad. Sci. U.S.A. 85:6409-6413(1988).
RN [5]
RP SEQUENCE OF 49-63; 197-204; 219-233 AND 261-275.
RC STRAIN=CD-1;
RX MEDLINE=93050795; PubMed=1330788;
RX Rosiere T.K., Wassarman P.M.;
RT "Identification of a region of mouse zona pellucida glycoprotein mZP3
RT that possesses sperm receptor activity."
RL Dev. Biol. 154:309-317(1992).
CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR

```

```

CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
CC SPECIES-SPECIFICITY OF THE INSEMINATION.
CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXTRACELLULAR
CC MATRIX.
CC -1- TISSUE SPECIFICITY: OOCYTES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE 2-WEEK GROWTH PHASE OF
CC OOGENESIS, PRIOR TO OVULATION.
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL: M20026; AAB18628.1; -.
CC EMBL: X14376; CAA32550.1; -.
CC PIR: S04189; S04189.
CC PIR: A30334; A30334.
CC PIR: A31232; A31232.
CC MGD: MGI:99215; ZP3.
CC InterPro: IPR001507; -.
CC Pfam: PF00100; zona_pellucida; 1.
CC PRINTS: PR00023; ZPELLUCIDA.
CC PROSITE: PS00682; ZP DOMAIN; 1.
CC Glycoprotein: Signal; Sulfatation; Sperm; Receptor; Transmembrane;
CC Extracellular matrix
CC SIGNAL 1 22 POTENTIAL.
CC CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
CC DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 388 408 POTENTIAL.
CC DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 45 308 ZP.
CC CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 424 AA: 46303 MW: 9089903FBD268365 CRC64;

Query Match 52.1%; Score 126; DB 1; Length 424;
Best Local Similarity 53.7%; Pred. No. 2,9e-08;
Matches 22; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 SWFPVGGPADICCCCKGCGTSPHSRRQPHVMSQMSRSVS 41
DB 309 SWFPVGGPADICCCCKGCGTSPHSRRQPHVMSQMSRSVS 349

RESULT 9
ZP3_PIG STANDARD; PRT; 421 AA.
ID ZP3_PIG
AC P42098;
ET 01-NOV-1995 (Rel. 32, Created)
ET 01-NOV-1995 (Rel. 32, Last sequence update)
ET 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3-BETA PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3-BETA) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
CN ZP3B OR ZPC.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RL Tissue=Ovary;
RA Yurewicz E.C., Hilder D., Fontenot G.K., Harris J.;

```

RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Ovary;  
 RA Okazaki Y., Sugimoto M.;  
 RU Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 CC  
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR  
 CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE  
 CC SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).  
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN  
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR  
 CC MATRIX.  
 CC -1- TISSUE SPECIFICITY: OOCYTES.  
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: L22169; AAA1145.1; -  
 CC EMBL: D45065; BAA08093.1; -  
 CC GLCOSULTEDB; P42098; -  
 CC InterPro: IPR001507; -  
 CC Pfam: PF00100; zona\_pellucida; 1.  
 CC PRINTS: PR00023; ZPELUCIDA.  
 CC PROSITE: PS00682; ZP\_DOMAIN, 1.  
 CC  
 CC GLYCOPROTEIN; Signal; Sulfatation; Sperm; Receptor; Transmembrane;  
 CC Extracellular matrix.  
 CC SIGNAL 1 21  
 CC CHAIN 1 22 421  
 CC FT DOMAIN 22 381 POTENTIAL.  
 CC FT TRANSMEM 382 402 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3-  
 CC FT DOMAIN 403 421 BETA.  
 CC FT DOMAIN 44 421 EXTRACELLULAR (POTENTIAL).  
 CC FT CARBOHYD 124 124 CYTOPLASMIC (POTENTIAL).  
 CC FT CARBOHYD 146 146 ZP.  
 CC FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CONFLICT 101 101 MISSING (IN REF. 2).  
 CC FT CONFLICT 107 107 D -> V (IN REF. 2).  
 CC FT CONFLICT 163 164 VF -> A (IN REF. 2).  
 CC FT CONFLICT 404 404 P -> A (IN REF. 2).  
 CC FT CONFLICT 404 404  
 CC SEQUENCE 421 AA; 46239 MW; DC97D760E985213A CRC64;

Query Match 49.28; Score 119; DB 1; Length 421;  
 Best Local Similarity 70.48; Pred. No. 2e-07;  
 Matches 19; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 WFPVGGPADICQCCNKGDCGTPSHSR 28  
 DB 308 WSPVEGPAVICRCCHKGCGTPSLSRK 334  
 ID ZP3\_BOVIN STANDARD; PRT; 421 AA.  
 AC P48830;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA  
 DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).  
 DE ZP3 OR ZPC.  
 OS Bos taurus (Bovine).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidea;  
 CC Bovidae; Bovinae; Bos.  
 CC NCBI\_TaxID:9913;  
 CC  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC TISSUE-Ovary;  
 CC TISSUE-Ovary;  
 CC MEDLINE:95143578; Pubmed-7841460;  
 CC Harris J.D., Hblier D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,  
 CC Sacco A.G.;  
 CC "Cloning and characterization of zona pellucida genes and cDNAs from  
 CC a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";  
 CC DNA Seq. 4:361-393(1994).  
 CC  
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR  
 CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE  
 CC SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).  
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN  
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR  
 CC MATRIX.  
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: U05775; AAA74385.1; -  
 CC InterPro: IPR001507; -  
 CC Pfam: PF00100; zona\_pellucida; 1.  
 CC PRINTS: PR00023; ZPELUCIDA.  
 CC PROSITE: PS00682; ZP\_DOMAIN, 1.  
 CC  
 CC GLYCOPROTEIN; Signal; Sulfatation; Sperm; Receptor; Transmembrane;  
 CC Extracellular matrix; Multigene family.  
 CC SIGNAL 1 22  
 CC CHAIN 1 23 421  
 CC FT DOMAIN 23 381 POTENTIAL.  
 CC FT TRANSMEM 382 402 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.  
 CC FT DOMAIN 403 421 EXTRACELLULAR (POTENTIAL).  
 CC FT CARBOHYD 124 124 CYTOPLASMIC (POTENTIAL).  
 CC FT CARBOHYD 146 146 ZP.  
 CC FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SEQUENCE 421 AA; 46545 MW; 905C4722B7B11DC CRC64;

Query Match 45.08; Score 109; DB 1; Length 421;  
 Best Local Similarity 56.78; Pred. No. 3.2e-06;  
 Matches 17; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 2 WFPVGGPADICQCCNKGDCGTPSHSR 31  
 DB 308 WSPVEGPTDICRCCKGRCGIGSRMRLSH 337  
 ID ZP3\_RABBIT STANDARD; PRT; 415 AA.  
 AC P48833;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA  
 DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C)  
 DE (FRAGMENT).  
 DE ZP3 OR ZPC.  
 OS Oryctolagus cuniculus (Rabbit).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
CC NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=95143578; PubMed=7841460;  
RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,  
RT "Cloning and characterization of zona pellucida genes and cDNAs from  
CC a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";  
CC DNA Seq. 4:361-393(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U05782; AAA74392.1; -  
DR InterPro; IPR001507; -  
DR Pfam; PF00100; zona\_pellucida; 1.  
DR PROSITE; PS00682; ZP\_DOMAIN; 1.  
KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;  
KW Extracellular matrix; Multigene family.  
FT NON\_TER 1  
FT SIGNAL <1 18 POTENTIAL.  
FT CHAIN 19 415 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.  
FT DOMAIN 19 378 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 379 399 POTENTIAL.  
FT DOMAIN 400 415 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 41 301 ZP.  
SQ SEQUENCE 415 AA; 44987 MW; 77396CF1BA3F5CB CRC64;  
Cy 1 SMFPVQGPADICCCCKNGDC---GTP-----SHSRROPHVMSO 35  
ID 302 SWAPVEGADICECCGNGCDLLAGSPMNQNHARSLSRRHVEE 348  
AC P32926; STANDARD; PRT; 999 AA.  
BT 01-OCT-1993 (Rel. 27, Created)  
ET 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE DESMOGLEIN 3 PRECURSOR (130 KDA PEMPHIGUS VULGARIS ANTIGEN) (PVA).  
EN DSG3.  
CS Homo sapiens (Human).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
CC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=92069753; PubMed=1720352;  
RA Anagat M., Klaus-Kovtun V., Stanley J.R.;  
RT "Autoantibodies against a novel epithelial cadherin in pemphigus  
CC vulgaris, a disease of cell adhesion.";  
CC Cell 67:869-877(1991).  
CC -1- FUNCTION: COMPONENT OF INTERCELLULAR DESMOsome JUNCTIONS.  
CC INVOLVED IN THE INTERACTION OF PLAQUE PROTEINS AND INTERMEDIATE  
CC FILAMENTS MEDIATING CELL-CELL ADHESION.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: EPIDERMIS, TONGUE, TONSIL, ESOPHAGUS AND  
CC CARCINOMAS.  
CC -1- DOMAIN: CALCIUM MAY BE BOUND BY THE CADHERIN-LIKE REPEATS

CC (POTENTIAL).  
CC -1- DISEASE: PEMPHIGUS VULGARIS (PV) IS A POTENTIALLY LETHAL SKIN  
CC DISEASE IN WHICH EPIDERMAL BLISTERS OCCUR AS THE RESULT OF THE  
CC LOSS OF CELL-CELL ADHESION CAUSED BY THE ACTION OF AUTOANTIBODIES  
CC AGAINST DSG3.  
CC -1- SIMILARITY: BELONGS TO THE CADHERIN FAMILY, DESMOsome SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M76482; AAA60230.1; -  
DR PIR; A41088; IJHUG3.  
DR HSSP; P09803; LEDH.  
DR MIM; 169615; -  
DR InterPro; IPR002126; -  
DR Pfam; PF00028; cadherin; 4.  
DR PROSITE; PS00232; CADHERIN\_1; 3.  
KW Cell adhesion; Signal; Transmembrane; Cytoskeleton; Glycoprotein;  
KW Calcium-binding; Repeat.  
FT SIGNAL 1 23 POTENTIAL.  
FT PROPEP 24 49 POTENTIAL.  
FT CHAIN 50 999 DESMOGLEIN 3.  
FT DOMAIN 50 615 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 616 640 POTENTIAL.  
FT DOMAIN 641 999 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 50 158 CADHERIN 1.  
FT DOMAIN 159 268 CADHERIN 2.  
FT DOMAIN 269 383 CADHERIN 3.  
FT DOMAIN 386 499 CADHERIN 4.  
FT REPEAT 910 935 DESMOGLEIN REPEAT 1.  
FT REPEAT 936 966 DESMOGLEIN REPEAT 2.  
FT CARBOHYD 110 110 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 180 180 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 459 459 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 545 545 N-LINKED (GLCNAc. . .) (POTENTIAL).  
SQ SEQUENCE 999 AA; 107503 MW; 60479DD46AC219A1 CRC64;  
Cy 1 DICOCCNKGDCGT-----PSHSRROPH 31  
ID 585 EVCCQDNNGICGTSTPTTSPGRTGRPH 612  
AC P31388; STANDARD; PRT; 436 AA.  
BT 01-JUL-1993 (Rel. 26, Created)  
ET 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE 5-HYDROXYTRYPTAMINE 6 RECEPTOR (5-HT-6) (SEROTONIN RECEPTOR)  
EN HTR6.  
CS Rattus norvegicus (Rat).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
CC NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Striatum;  
RX Monma F.J. Jr., Shen Y., Ward R.P., Hamblin M.W., Sibley D.R.;  
RT "Cloning and expression of a novel serotonin receptor with high  
CC affinity for tricyclic psychotropic drugs.";

RL	Mol. Pharmacol. 43:320-327(1993).		
RN	SEQUENCE FROM N.A.		
RX	MEDLINE=93277562; PubMed=8389146;		
RA	Ruat M., Traiffort E., Arrang J.-M., Tardivel-Lacombe J., Diaz J.,		
RA	Leus R., Schwartz J.-C.;		
RT	"A novel rat serotonin (5-HT6) receptor: molecular cloning,		
RT	localization and stimulation of cAMP accumulation.";		
RL	Biochem. Biophys. Res. Commun. 193:268-276(1993).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-WISTAR;		
RA	Marital R.;		
CC	Submitted (xxx-1993) to the EMBL/Genbank/DBJ databases.		
CC	-I- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR		
CC	5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTION		
CC	AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF		
CC	THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT STIMULATES ADENYLATE		
CC	CYCLASE. IT HAS A HIGH AFFINITY FOR TRICYCLIC PSYCHOTROPIC DRUGS.		
CC	-I- SUBCELLULAR LOCATION: INTERNAL MEMBRANE PROTEIN.		
CC	-I- TISSUE SPECIFICITY: LOCALIZED EXCLUSIVELY IN THE CENTRAL NERVOUS		
CC	SYSTEM. PREDOMINANTLY IN THE CORPUS STRIATUM BUT ALSO IN VARIOUS		
CC	LYMBIC AND CORTICAL REGIONS.		
CC	-I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see <a href="http://www.isb-slb.ch/announce/">http://www.isb-slb.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).		
CC	-----		
DR	EMBL; L03202; AAA40618.1; -		
DR	EMBL; S62043; AAB26908.1; -		
DR	EMBL; L19656; AAA40611.1; -		
DR	GCRDB; GCR_0723; -		
DR	GCRDB; GCR_0719; -		
DR	InterPro; IPR000276; -		
DR	InterPro; IPR002232; -		
DR	Pfam; PF00001; 7tm_1; 1.		
DR	PRINTS; PR0237; GPCRHDOPSN.		
DR	PRINTS; PR01102; 5HT6RECEPT.		
DR	PROSITE; PS00237; G_PROTEIN_REC.F1.1; 1.		
DR	PROSITE; PS50262; G_PROTEIN_REC.F1.2; 1.		
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;		
KW	MultiGene family.		
FT	DOMAIN	1	34
FT	TRANSMEM	35	57
FT	DOMAIN	58	64
FT	TRANSMEM	65	85
FT	DOMAIN	86	100
FT	TRANSMEM	101	122
FT	DOMAIN	123	144
FT	TRANSMEM	145	166
FT	DOMAIN	167	184
FT	TRANSMEM	185	208
FT	DOMAIN	209	265
FT	TRANSMEM	266	290
FT	DOMAIN	291	295
FT	TRANSMEM	296	320
FT	DOMAIN	321	436
FT	CARBOHYD	9	9
FT	DISULFID	99	180
FT	CONFLICT	57	57
FT	CONFLICT	336	436
FT	SEQUENCE	436 AA;	46922 MW;
FT			68C00F6A3CB5FB4 CRC64;

```

Query Match      22.9% Score 55.5; DB 1; Length 436;
Best Local Similarity 34.9%; Pred. No. 9.4;
Matches 15; Conservative 3; Mismatches 16; Indels 9; Gaps 2;

Oy      4 PVOGPAADICQ-----CGNKGDGCTPSSHRSRQPHVMSOKSRVS 41
       1:| | | | | | | | | | | | | | | | | | | | | |
Db      382 PLUGAPRACSCSPSFCLEPRPPGTFRHPPGP-----LMSTSLIS 420

RESULT 14
PML1_HUMAN
ID      PML1_HUMAN          STANDARD:          PRT:          865 AA.
CD      O43490.
DT      01-OCT-2000 (Rel. 40, Created)
DT      01-OCT-2000 (Rel. 40, Last sequence update)
DT      01-OCT-2000 (Rel. 40, Last annotation update)
DE      PROMININ-LIKE PROTEIN 1 PRECURSOR (ANTIGEN AC133).
GN      PROML1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RC      SEQUENCE FROM N.A., AND SEQUENCE OF 31-42; 200-211; 280-291 & 641-656.
RP      TISSUE=Fetal liver, AND Retinoblastoma;
RX      MEDLINE=98052559; PubMed=9389721;
RA      Miraglia S., Godfrey W., Yin A.H., Atkins K., Warnke R., Holden J.T.,
RA      Bray R.A., Waller E.K., Buck D.W.;
RT      "A novel five-transmembrane hematopoietic stem cell antigen:
RT      isolation, characterization, and molecular cloning.";
RL      Blood 90:5013-5021(1997).
RN      [2]
RC      CHARACTERIZATION.
RP      TISSUE=Fetal liver;
RX      MEDLINE=98052558; PubMed=9389720;
RA      Yin A.H., Miraglia S., Zanjani E.D., Almeida-Porada G., Ogawa M.,
RA      Leary A.G., Olweiss J., Kearney J., Buck D.W.;
RT      "AC133, a novel marker for human hematopoietic stem and progenitor
RT      cells.";
RL      Blood 90:5002-5012(1997).
RN      [3]
RC      DISease.
RX      MEDLINE=20056119; PubMed=10587575;
RA      May M.A., Corbelli D., Koch J., Hellwig A., Wilson-Wheeler J.C.,
RA      Bridges R.J., Kumarmanickavel G., John S., Nancarrow D., Roeper K.,
RA      Weigmann A., Hutter W.B., Denton M.J.;
RT      "A frameshift mutation in prominin (mouse)-like 1 causes human retinal
RT      degeneration.";
RL      Hum. Mol. Genet. 9:27-34(2000).
CC      -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC      -I- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED ON CD34 HEMATOPOIETIC
CC      STEM AND PROGENITOR CELLS IN ADULT AND FETAL BONE MARROW, PETAL
CC      LAYER, COND BLOOD AND ADULT PERIPHERAL BLOOD. NOT DETECTED ON
CC      OTHER BLOOD CELLS. ALSO EXPRESSED IN A NUMBER OF NONLYMPHOID
CC      TISSUES INCLUDING RETINA, PANCREAS, PLACENTA, KIDNEY, LIVER, LUNG,
CC      BRAIN AND HEART.
CC      -I- PIN: GLYCOSYLATED.
CC      -I- DISEASE: DEFECTS IN PROML1 ARE THE CAUSE OF AN AUTOSOMAL RECESSIVE
CC      FORM OF RETINAL DEGENERATION CHARACTERIZED BY NIGHT BLINDNESS AND
CC      LOSS OF PERIPHERAL VISION FROM CHILDHOOD WITH PROGRESSION TO
CC      PROFOUND VISUAL IMPAIRMENT AND EXTINGUISHED ELECTRORETINOGRAMS BY
CC      THE THIRD DECADE. A SINGLE NUCLEOTIDE DELETION YIELDS A TRUNCATED
CC      PROTEIN LACKING HALF OF THE SECOND EXTRACELLULAR LOOP, THE FINAL
CC      MEMBRANE-SPANNING DOMAIN AND THE CYTOPLASMIC C-TERMINUS.
CC      -I- SIMILARITY: BELONGS TO THE PROMININ FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC      the European Bioinformatics institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by web and for commercial
CC      entities requires a license agreement (see http://www.isb-sdb.ch/announcement/
```

```

C3 or send an email to license@lsb.sib.ch).
C3 -----
C3 EMBL; AF027208; AAB92514.1; -.
C3 DR MIM; 604365; -.
C3 KW Transmembrane; glycoprotein; Antigen; Signal; Polymorphism;
C3 Vision.
C3 FT SIGNAL 1 19 POTENTIAL.
C3 FT CHAIN 20 865 PROMININ-LIKE PROTEIN 1.
C3 FT DOMAIN 20 108 EXTRACELLULAR (POTENTIAL).
C3 FT TRANSMEM 109 129 POTENTIAL.
C3 FT DOMAIN 130 157 CYTOPLASMIC (POTENTIAL).
C3 FT TRANSMEM 158 178 POTENTIAL.
C3 FT DOMAIN 179 433 EXTRACELLULAR (POTENTIAL).
C3 FT TRANSMEM 434 454 POTENTIAL.
C3 FT DOMAIN 455 486 CYTOPLASMIC (POTENTIAL).
C3 FT TRANSMEM 487 507 POTENTIAL.
C3 FT DOMAIN 508 792 EXTRACELLULAR (POTENTIAL).
C3 FT TRANSMEM 793 813 POTENTIAL.
C3 FT DOMAIN 814 865 CYTOPLASMIC (POTENTIAL).
C3 FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
C3 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
C3 FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
C3 FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
C3 FT CARBOHYD 548 548 N-LINKED (GLCNAC. . .) (POTENTIAL).
C3 FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
C3 FT CARBOHYD 729 729 N-LINKED (GLCNAC. . .) (POTENTIAL).
C3 FT CARBOHYD 730 730 N-LINKED (GLCNAC. . .) (POTENTIAL).
C3 FT VARIANT 31 31 A -> G.
C3 FT VARIANT /FTID-VAR_010382.
C3 FT VARIANT A -> S.
C3 FT VARIANT /FTID-VAR_010383.
C3 FT CONFLICT 200 200 D -> P OR A (IN AA SEQUENCE).
C3 FT CONFLICT 284 284 S -> D (IN AA SEQUENCE).
C3 FT CONFLICT 288 288 S -> R (IN AA SEQUENCE).
C3 SC SEQUENCE 865 AA; 97201 MW; D21CB05AD3ADEDF CRC64;

Query Match 22.7%; Score 55; DB 1; Length 865;
Best Local Similarity 47.4%; Pred. No. 19;
Matches 9; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 11 ICOCCKNGDCGTPSHSRRO 29
Db 132 MCRCCNK--CGEMHQRK 148

RESULT 15
PROM_MOUSE STANDARD; PRT; 867 AA.
ID PROM_MOUSE
AC 034990; 035408;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROMININ PRECURSOR (ANTIGEN AC133 HOMOLOG).
EN PROM.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RX NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RC Miragila S., Godfrey W., Buck D.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RC MEDLINE=98024147; PubMed=9356465;
RX Weigmann A., Corbelli D., Hellwig A., Huttner W.B.;
RT "Prominin A, a novel microvilli-specific polytopic membrane protein of
RT the apical surface of epithelial cells, is targeted to plasmalemma
RT protrusions of non-epithelial cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:12425-12430(1997).

```

CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).	
CC	-1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED ON THE APICAL SIDE OF	
CC	NEUROEPITHELIAL CELLS AND OF OTHER EPITHELIA SUCH AS LUNG BUDS,	
CC	GUT AND URETERH BUDS. IN THE ADULT, EXPRESSED AT THE APICAL SIDE	
CC	OF THE KIDNEY TUBULES AND OF THE EPENDYMAL LAYER OF THE BRAIN. NOT	
CC	EXPRESSED IN GUT, LIVER, LUNG, PITUITARY, ADRENAL, HEART OR	
CC	SPLEEN.	
CC	-1- SIMILARITY: BELONGS TO THE PROMININ FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; AF039663; AAB6916.1; -	
DR	EMBL; AF026269; AAB6715.1; -	
DR	MGI:1100886; Prom.	
KM	Signal; Transmembrane; Glycoprotein.	
FT	SIGNAL	1
FT	CHAIN	19
FT	PROMININ.	
FT	DOMAIN	20
FT	TRANSMEM	107
FT	DOMAIN	108
FT	TRANSMEM	128
FT	DOMAIN	129
FT	TRANSMEM	158
FT	DOMAIN	159
FT	TRANSMEM	179
FT	DOMAIN	180
FT	TRANSMEM	434
FT	DOMAIN	435
FT	TRANSMEM	435
FT	DOMAIN	455
FT	TRANSMEM	456
FT	DOMAIN	487
FT	TRANSMEM	488
FT	DOMAIN	509
FT	TRANSMEM	794
FT	DOMAIN	795
FT	TRANSMEM	815
FT	DOMAIN	816
FT	CARBOHYD	867
FT	CARBOHYD	273
FT	CARBOHYD	273
FT	CARBOHYD	291
FT	CARBOHYD	291
FT	CARBOHYD	332
FT	CARBOHYD	332
FT	CARBOHYD	374
FT	CARBOHYD	374
FT	CARBOHYD	415
FT	CARBOHYD	415
FT	CARBOHYD	554
FT	CARBOHYD	554
FT	CARBOHYD	581
FT	CARBOHYD	581
FT	CARBOHYD	732
FT	CARBOHYD	732
FT	CONFLICT	64
FT	CONFLICT	64
FT	CONFLICT	84
FT	CONFLICT	84
FT	CONFLICT	102
FT	CONFLICT	102
FT	CONFLICT	668
FT	CONFLICT	668
FT	CONFLICT	844
FT	CONFLICT	844
SQ	SEQUENCE	867 AA; 97112 MW; D442F6372552B3C8 CRC64;

Search completed: July 12, 2001, 08:27:09  
Job time: 179 sec

```

Query Match          22.7%; Score 55; DB 1; Length 867;
Best Local Similarity 47.4%; Pred. No. 19;
Matches 9; Conservative 4; Mismatches 4; Indels 2; Gaps 1;
QY 11 ICQCCNKGDCTPPSHSRQ 29
      :|:|:|:|  |  |  |:|:|
DB 133 MCRCCNK--CGGEMHQRQK 149

```

Search completed: July 12, 2001, 08:27:09  
Job time: 179 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

ON protein - protein search, using sw model

Run on: July 12, 2001, 08:24:10 ; Search time 21.08 Seconds  
(without alignments)  
257.330 Million cell updates/sec

Title: US-09-252-828a-2

Sequence: 1 SWFVQGPADICCCCKGDC.....TPSHSRPHVMSQMSRSVS 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

SPTREMBL.16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	217	89.7	210	4 Q12903	homo sapien
2	125	51.7	424	11 P97708	rattus norv
3	115	47.5	418	11 Q9ER10	microtus br
4	100.5	41.5	454	13 Q91675	xenopus lae
5	100.5	41.5	460	13 Q91728	xenopus lae
6	98.5	40.7	446	13 P79762	gallus gall
7	98.5	40.7	446	13 P9PWF8	gallus gall
8	95	39.3	422	6 Q77685	trichosurus
9	87	36.0	446	13 Q73670	coturnix co
10	60.5	25.0	99	2 Q9PE89	xyella fas
11	60.5	25.0	436	13 Q9DG37	strepomyce
12	59.5	24.6	934	2 Q05478	strepomyce
13	57	23.6	1130	5 Q20778	caenorhabd
14	56.5	23.3	819	2 Q9PK14	chlamydia m
15	56	23.1	122	4 Q9P166	homo sapien
16	56	23.1	448	10 Q9FYV3	saccharum o
17	56	23.1	641	11 Q08463	rattus norv
18	56	23.1	642	11 Q70421	mus musculu
19	56	23.1	647	4 Q94815	homo sapien

20	56	23.1	648	4 Q9UP38	homo sapien
21	55.5	22.9	344	11 Q54788	mus musculu
22	55.5	22.9	435	13 Q9PU06	carassius a
23	55.5	22.9	447	10 Q9ZRP9	malus domes
24	55	22.7	295	2 Q44588	alcaligenes
25	55	22.7	591	11 Q9J149	rattus norv
26	54.5	22.5	340	14 Q9JF36	vacinia vi
27	54.5	22.5	424	13 Q92027	cyprinus ca
28	54.5	22.5	559	13 Q919M5	cyprinus lae
29	54.5	22.5	735	11 Q60473	cavia porce
30	54	22.3	214	2 Q9RD56	strepomyce
31	54	22.3	214	6 Q97602	oryctolagus
32	54	22.3	315	6 Q97601	oryctolagus
33	54	22.3	477	6 Q97600	oryctolagus
34	54	22.3	537	4 Q9HBE2	homo sapien
35	54	22.3	544	1 Q48304	haloterax s
36	54	22.3	592	13 Q57328	gallus gall
37	54	22.3	609	4 Q9HD61	homo sapien
38	53.5	22.1	237	5 Q9VAG7	dtrosophila
39	53.5	22.1	338	4 Q60521	homo sapien
40	53.5	22.1	351	1 Q9YA98	aeropyrum p
41	53.5	22.1	379	10 Q9EYF9	arabidopsis
42	53.5	22.1	738	4 Q60327	homo sapien
43	53.5	22.1	2168	5 Q19330	caenorhabd
44	53	21.9	197	6 Q28584	ovis aries
45	53	21.9	226	10 Q9SBR2	fragaria an

## ALIGNMENTS

RESULT 1  
Q12903 PRELIMINARY; PRT; 210 AA.  
ID Q12903;  
AC Q12903;  
DT 01-NOV-1996 (TREMUREL. 01, Created)  
DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)  
DT 01-MAY-2000 (TREMUREL. 13, Last annotation update)  
DE POM-2P3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=OVARY;  
RX MEDLINE=95309900; Pubmed=7789967;  
RA Kiersztoz S., Osawa G.A., Liang L.F., Modi W.S., Dean J.;  
RT "POM-2P3, a bipartite transcript derived from human ZP3 and a POM121  
homologue."  
RL Genomics 25:354-359(1995).  
DR EMBL; U10099; AAA85788.1; -.  
DR InterPro: IPR001507; -.  
DR Pfam: PF00100; zona\_pellucida; 1.  
SQ SEQUENCE 210 AA; 23196 MM; 6E9723ECLB77C8D0 CRC64;

Query Match 89.7%; Score 217; DB 4; Length 210;  
Best Local Similarity 87.8%; Pred. No. 3.6e-22;

Matches 36; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SWFVQGPADICCCCKGDCGTPSHSRPHVMSQMSRSVS 41  
DB 146 SWFVEGPADICCCCKGDCGTPSHSRPHVMSQMSRSVS 186

RESULT 2  
ID P97708 PRELIMINARY; PRT; 424 AA.  
AC P97708; O55084;  
DT 01-NOV-1998 (TREMUREL. 08, Created)  
DT 01-NOV-1998 (TREMUREL. 08, Last sequence update)  
DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)

DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR  
 DE (ZONA PELLUCIDA GLYCOPROTEIN ZP3) (ZONA PELLUCIDA GLYCOPROTEIN 3).  
 DE ZP3 OR ZP-3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WISTAR; TISSUE-OVARY;  
 RA Macduff P.E., Kerr L.E., Aitken R.J.;  
 RU J. Reprod. Fert. 11. Abstr. Ser. 18:86-86(1996).  
 RU [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-OVARY;  
 RA Akatsuka K., Yoshida-Komiya H., Tulsiani D.P., Orgebin-Crist M.,  
 RA Hiroi M., Arai Y.;  
 RU Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RU [1]  
 CC FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR  
 CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE  
 CC SPECIES-SPECIFICITY OF THE INSEMINATION.  
 CC -1 SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH  
 CC ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.  
 CC -1 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR  
 CC MATRIX.  
 CC -1 PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY  
 CC SIMILARITY).  
 CC -1 SIMILARITY: CONTAINS A ZP DOMAIN, WHICH CURRENTLY HAS BEEN FOUND  
 CC IN ZP2, ZP3, GP2, TGF-3 AND UROMODULIN.  
 CC EMBL: Y10823; CAA71787.1; -;  
 CC EMBL: D78482; BA024456.1; -;  
 CC InterPro: IPR001507; -;  
 CC Pfam: PF00100; zona\_pellucida; 1.  
 DR PRINTS: PR00023; zpellucida.  
 DR PROSITE: PS00682; ZP DOMAIN; FALSE\_NEG.  
 DR SMART: SM00241; ZP. 1.  
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;  
 KW Extracellular matrix.  
 FT SIGNAL 1 16 POTENTIAL.  
 FT CHAIN 17 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.  
 FT DOMAIN 17 387 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 388 408 POTENTIAL.  
 FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 45 308 ZP.  
 FT DOMAIN 329 334 POLY-SER.  
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 355 355 V -> A (IN REF. 2).  
 FT CONFLICT 112 112 N -> S (IN REF. 2).  
 FT CONFLICT 412 412 K -> M (IN REF. 2).  
 SQ SEQUENCE 424 AA; 45900 MW; 2AB42CBB14DE8701 CRC64;

Query Match 51.7%; Score 125; DB 11; Length 424;  
 Best Local Similarity 51.2%; Pred. No. 2,6e-09;  
 Matches 21; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 SWFVPGADICCCCKNGDCGTPSHSRQPHVMSQWSRSVS 41  
 DB 309 SWLPVEGDADICCCCKNGDCGTPSHSRQPHVMSQWSRSVS 349

RESULT 3  
 Q9ER10 PRELIMINARY; PRT; 418 AA.  
 AC Q9ER10;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE ZONA PELLUCIDA 3 GLYCOPROTEIN.

OS Microtus brandti.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Arvicolinae;  
 OC Microtus.  
 OC NCBI\_TaxID=137787;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OVARY;  
 RA Li H., Piao Y.S., Zhang Z.B., Zhu C., Hardy C., Hinds L.A.;  
 RU Cloning and characterization of zona pellucida cDNA from Brandt's  
 RU vole " "  
 RU Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF304487; AAG18455.1; -;  
 SQ SEQUENCE 418 AA; 45662 MW; 12C05062555F468F CRC64;

Query Match 47.5%; Score 115; DB 11; Length 418;  
 Best Local Similarity 44.7%; Pred. No. 6e-08;  
 Matches 17; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 SWFVPGADICCCCKNGDCGTPSHSRQPHVMSQWSR 38  
 DB 309 SWLPVEGDTPVCCCTGDCSSSRYSRRAHVAVPRNR 346

RESULT 4  
 Q91675 PRELIMINARY; PRT; 454 AA.  
 AC Q91675;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE ZONA PELLUCIDA C GLYCOPROTEIN PRECURSOR.  
 GN XLZPC.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hedrick J.L., Yang J.C.;  
 RU Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U44952; AAB39079.1; -;  
 DR InterPro: IPR001507; -;  
 DR Pfam: PF00100; zona\_pellucida; 1.  
 DR SMART: SM00241; ZP. 1.  
 KW Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 454 ZONA PELLUCIDA C GLYCOPROTEIN.  
 SQ SEQUENCE 454 AA; 49589 MW; A8DB9A5B978B5197 CRC64;

Query Match 41.5%; Score 100.5; DB 13; Length 454;  
 Best Local Similarity 46.3%; Pred. No. 6,2e-06;  
 Matches 19; Conservative 9; Mismatches 10; Indels 3; Gaps 3;

QY 1 SWFVPGADICCCCKNGDCGTPSHSRQPHVMSQWSR 38  
 DB 330 SWSPDQSPNISCDCGTGNCVYFGQSRRLGPIYSGSRMNO 370

RESULT 5  
 Q91728 PRELIMINARY; PRT; 460 AA.  
 AC Q91728;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE GP43.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;

	DE	ZONA PELLUCCIDA C PROTEIN.	
	OS	Gallus gallus (Chicken).	
	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	OC	Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	
	OC	Gallus.	
	OX	NCBI_TaxID=9031;	
	RN	[1]	
	RN	SEQUENCE FROM N.A.	
	RP	TTSUDE-LIVER;	
	RC	Kono Y., Matsuda T.;	
	RT	"gallus gallus zona pellucida C protein gene.";	
	RL	Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.	
	DR	EMBL; AB031033; BAA83418.1; -.	
	DR	InterPro: IPR001179; -.	
	DR	InterPro: IPR001507; -.	
	DR	Pfam: PF00100; zona_pellucida; 1.	
	DR	PRINTS: PR00023; ZPELLUCIDA.	
	DR	PROSITE: PS00453; FKBP_PPIASE_1; UNKNOWN_1.	
	DR	SMART: SM00241; ZP; 1.	
	SQ	SEQUENCE 446 AA; 47670 MW; DC5409532C22F03C CRC64;	
	QY	1 SWPVGGPADICCCCKNGDCCTPSHSRRQPHVMSGW 36	
	Db	: : : : : : : : : : : : : : : : : : :	
		330 TWVEGSRDYVCNCCEGTGCNECPALSRRL-NPMERW 364	
	RESULT	8	
	077685	PRELIMINARY:	PRT: 422 AA.
	ID	077685	
	AC	077685;	
	DT	01-NOV-1998 (TREMBlrel. 08, Created)	
	DT	01-NOV-1998 (TREMBlrel. 08, Last sequence update)	
	DE	01-MAR-2001 (TREMBlrel. 16, Last annotation update)	
	DE	ZONA PELLUCCIDA 3 PROTEIN.	
	GN	ZP3.	
	OS	Trichosurus vulpecula (Brush-tailed possum).	
	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	OC	Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.	
	OX	NCBI_TaxID=9337;	
	RN	[1]	
	RN	SEQUENCE FROM N.A.	
	RP	McCartney C.A., Mate K.E.;	
	RA	"Cloning and characterization of a zona pellucida 3 cDNA from a	
	RT	marsupial, the brush-tail possum Trichosurus vulpecula.";	
	RL	Zygotc 0.10-0.1(1998).	
	DR	EMBL; AF079524; AAC28736.1; -.	
	DR	InterPro: IPR001179; -.	
	DR	InterPro: IPR001507; -.	
	DR	Pfam: PF00100; zona_pellucida; 1.	
	DR	PRINTS: PR00023; ZPELLUCIDA.	
	DR	PROSITE: PS00453; FKBP_PPIASE_1; UNKNOWN_1.	
	DR	SMART: SM00241; ZP; 1.	
	SQ	SEQUENCE 422 AA; 45714 MW; 58BA61A92F82612A CRC64;	
	Query Match	39.3%; Score 95; DB 6; Length 422;	
	Best Local Similarity	53.6%; Pred. No. 3.3e-05;	
	Matches 15; Conservative	4; Mismatches 9; Indels 0; Gaps 0;	
	QY	2 WFPVGADICCCCKNGDCGTPSHRSRRO 29	
	Db	: : : : : : : : : : : : : : : : : : :	
		314 WLVEGPRDICSCTGTCTISLSRRK 341	
	RESULT	9	
	073670	PRELIMINARY:	PRT: 446 AA.
	ID	073670	
	AC	073670;	

DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE OVALI ZPC.  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix.  
 NCBI\_TaxID=93934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RL Kono Y., Tsuda T., Aoki N., Kitajima K., Matsuda T.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB012606; BAA25637.1;  
 DR InterPro: IPR001179;  
 DR InterPro: IPR001507;  
 DR Pfam: PF00100; zona\_peptidase\_1.  
 DR PRINTS; PR00023; ZPELUCIDA.  
 DR PROSITE; PS00453; FKBP\_PPIASE\_1; UNKNOWN\_1.  
 DR SMART; SM00241; ZP; 1.  
 SQ SEQUENCE 446 AA; 47621 MW; 67A98797A58C135B CRC64;

Query Match 36.0%; Score 87; DB 13; Length 446;  
 Best Local Similarity 37.8%; Pred. No. 0.00043;  
 Matches 14; Conservative 6; Mismatches 13; Indels 4; Gaps 1;

QY 1 SWFVPGADICCCNKGDCGTPSHSRQPHVMSQWS 37  
 DB 330 TWVPEGSRDVCSCEFGNCAPALRR-----LNNWS 362

RESULT 10  
 Q9PE89 PRELIMINARY; PRT; 99 AA.  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL PROTEIN XF1139.  
 GN XF1139.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9A5C.  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
 RA Faciniani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,  
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kiriama E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marinho C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matukuma A.Y.,  
 RA Menck C.F.M., Miranda E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi W.F., Trufi D., Tsai S.M., Tsunako M.H.,

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."  
 RL Nature 406:151-159(2000).  
 DR EMBL; AE003949; AAF83949.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 99 AA; 11303 MW; 2CF720F47433EF81 CRC64;

Query Match 25.0%; Score 60.5; DB 2; Length 99;  
 Best Local Similarity 45.8%; Pred. No. 0.49;  
 Matches 11; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 12 CQCNKGDCGTPSHSRQPHVMSQ 35  
 DB 36 CRCARK-HCGLPEHSRTPFNPEQ 58

RESULT 11  
 Q9DG37 PRELIMINARY; PRT; 436 AA.  
 AC Q9DG37;

DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ZONA PEPTIDICA GLYCOPROTEIN 3 (FRAGMENT).  
 GN ZP3.  
 OS Pimephales promelas.  
 OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Cyprinidae; Leuciscinae; Pimephales.  
 NCBI\_TaxID=90988;

RP SEQUENCE FROM N.A.  
 RA Morris C.A., Thomas-Jones R.E., Cryer J., Woodhead S., Kille P.;  
 RT "Identification of the ZP3 cDNA sequence from *Pimephales promelas*  
 RT exposed to diethylstilbestrol.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF192407; AAG28398.1;  
 FT NON\_TER 1  
 SQ SEQUENCE 436 AA; 47690 MW; 8C54DC948DBC41B6 CRC64;

Query Match 25.0%; Score 60.5; DB 13; Length 436;  
 Best Local Similarity 27.5%; Pred. No. 1.8;  
 Matches 11; Conservative 4; Mismatches 24; Indels 1; Gaps 1;

QY 2 WFPVPGADICCCNKGDCGTPSHSRQPHVMSQWSRSVS 41  
 DB 385 WFAADGNDVCGCCD-STCGPDGEFASPYGGIWECKAS 423

RESULT 12  
 ID 005478 PRELIMINARY; PRT; 934 AA.  
 AC 005478;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE LIPA TRANSCRIPTIONAL ACTIVATOR.  
 GN LIPR.  
 OS Streptomyces sp.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 NCBI\_TaxID=1931;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M11;  
 RX MEDLINE=93138418; PubMed=8422993;  
 RA Perez C., Juarez K., Garcia-Castells E., Soberon G.,  
 RA Servin-Sanzalez L.;  
 RT "Cloning, characterization, and expression in *Streptomyces lividans* 66  
 RT of an extracellular lipase-encoding gene from *Streptomyces* sp. M11.";



